Unsupervised Short-term Covariate Shift Minimization for Self-paced BCI

Mohammadi, R., Mahloojifar, A., & Coyle, D. (2013). Unsupervised Short-term Covariate Shift Minimization for Self-paced BCI. In *Unknown Host Publication* (pp. 101-106). IEEE.

Published in:
Unknown Host Publication

Publication Status:
Published (in print/issue): 01/01/2013

Document Version
Publisher's PDF, also known as Version of record

Unsupervised Short-term Covariate Shift Minimization for Self-paced BCI

Raheleh Mohammadi, Ali Mahloojifar
Biomedical Engineering Department
Tarbiat Modares University
Tehran, Iran
(raheleh.mohammadi & mahlooji)@modares.ac.ir

Damien Coyle
Intelligent Systems Research Center
University of Ulster,
Derry, UK
dh.coyle@ulster.ac.uk

***Abstract*—A major challenge for Brain Computer Interface systems (BCIs) is dealing with non-stationarity in the EEG signal. There are two types of EEG non-stationarities 1) long-term changes related to fatigue, changes in recording conditions or effects of feedback training which is addressed in classification step and 2) short-term changes related to different mental activities and drifts in slow cortical potentials which can be addressed in the feature extraction step. In this paper we use a covariate shift minimization (CSM) method to alleviate the short-term (single trial) effects of EEG non-stationarity to improve the performance of self-paced BCIs in detecting foot movement from the continuous EEG signal. The results of applying this unsupervised covariate shift minimization with two different classifiers, linear discriminant analysis (LDA) and probabilistic classification vector machines (PCVMs) along with two different filtering methods (constant bandwidth and constant-Q filters) show the considerable improvement in system performance.**

Keywords—Covariate shift minimization; EEG; Non-stationarity; Self-paced Brain Computer interfaces

I. Introduction

The goal of a brain-computer interface (BCI) is to infer a user's intention by translating their brain pattern into a controlling signal for different applications. This system is especially designed to assist people suffering from neurological disorders [1]. Electroencephalogram (EEG) signals recorded non-invasively and deployed in BCI are inherently non-stationary resulting in substantial change over time, both within a single session and between sessions, resulting in significant challenges in maintaining BCI system robustness [2].

There are two types of non-stationarities: short-term changes related to different mental activities (e.g. hand movement, mental arithmetic, slow cortical potentials etc.), and long term changes related to fatigue, small differences in electrode position, or effects of feedback training [3]. To build a stable BCI, the system should be able to adapt to these non-stationary changes in the EEG signal. The data processing in BCIs consists typically of two main steps, (i) signal processing and feature extraction, and (ii) classification or feature translation. Adaptation can be done on one of the above steps. The short-term drifts are normally addressed in the feature extraction step while the second type of non-stationarities or long-term changes are normally addressed in the classification and feature translation step [3].

Several studies have shown the usefulness of applying adaptive methods in enhancing the system performance. The adaptive BCIs described in most of the publications are categorized in two different groups; supervised and unsupervised adaptive systems [4]. Supervised adaptation is a situation that the true class labels are known in advance while unsupervised adaptation is based on unlabeled data. Unsupervised classifier adaptation has received more attention in the literature [4-8] while the number of systems designed with an adaptive feature extraction block, are very limited. The unsupervised covariate shift minimization (CSM) method proposed in [9] is a feature adaptation method which alleviates the non-stationary effects between sessions in a 2 class motor imagery based BCI. CSM involves estimating feature deviation from the mean over time using polynomials and subsequently accounting for the drift by subtracting the estimated drift. The term covariate shift refer to the case in which the training and testing feature distributions change while the conditional distribution of the classifier output given an input is unchanged. Based on this definition, Sugiyama et. al. [10] proposed a modification of the cross-validation technique called importance weighted cross-validation (IWCV) that can be used for model and parameter selection in classification tasks. In [11] the covariate shift adaptation method proposed using IWCV to select the parameters of an importance weighted LDA (IWLDA). While Covariate shift adaptation method [11] tries to adapt the parameters of the classifier, in CSM [9] the adaptation is done in feature extraction step.

In this paper we show that by applying the CSM method in a foot movement based self-paced BCI we can minimize the short-term non-stationary effects of the signal in a single session and improve performance of the system in detecting foot movement in the continuous EEG signal. Although the CSM method has been demonstrated in [9] using only one-feature adaptation and later using multiple feature adaptation (in press) to improve inter-session unsupervised adaptation, this work presents intra-session short term multiple feature adaptation in self-paced mode for the first time. The problem of non-stationarity of EEG data in self-paced systems have been addressed in classification [12, 13] and post-processing

978-1-4673-5871-2/13/$31.00

[14] steps before. An unsupervised adaptive method based on sequential expectation maximization for Gaussian mixture model (GMM) proposed in [12] for onset detection problem. In [12] the mean and variance of each Gaussian component are adapted online while in [13] the number of Gaussian components in GMM was adapted online. In another study Tsui et al. [14] adapted the refractory and dwell windows required for post processing in self-paced BCI control.

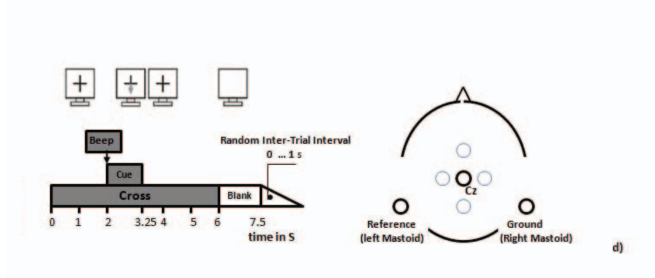
The remainder of the paper is organized into four sections. Section II outlines data acquisition. Section III contains the details of the methodology for frequency decomposition, feature extraction, covariate shift minimization, classification and post-processing. Results and discussion are presented in section IV. Finally conclusions are presented in section V.

II. Data Description

The analysis is performed on data provided by the laboratory of Brain Computer Interfaces (BCI-Lab), Graz University of Technology [15]. Seven healthy subjects participated in the study. Each subject performed 3 runs (each run comprised 30 trials). The subjects performed a brisk movement of both feet after the presentation of the cue. At the beginning of the trial (t=0) a “+” was presented; then at t=2 the presentation of an arrow pointing downwards cued the subject to perform a brisk foot movement of both feet for about 1 second duration. The cross and cue disappear at t=3.25s and at t=6s, respectively. At t=7.5 the trial ends (Fig. 1(a)). The sampling frequency was 250 Hz. Our analysis is performed on a single small Laplacian derivation over the Cz electrode (Fig. 1(b)).

III. Method

The single EEG channel derived from 1st and 2nd run of each subject is used as training data and 3rd run is used as test data. A block diagram of the proposed system is illustrated in Fig. 2.



(a) (b)

Figure 1. Paradigm and electrode positions a) timing scheme of each trial, b)Electrode position

One-second sliding windows of EEG signal with 0.8 seconds overlap are first filtered in a broad frequency range of 6 to 36 Hz which covers the mu and beta rhythms. The logarithmic band power of the signals then calculated to extract features in different frequency bands and finally the classifier is trained to reveal the event-related synchronization of the EEG signal which occurs after foot movement. Since the training data and test data features follow different distributions, the classifiers trained on training data are not optimum for the test data. Therefore in the test phase, the shifts of the features are estimated and removed before classification using the covariate shift minimization (CSM) approach. In order to examine the effectiveness of the proposed CSM method regardless of the feature extraction and classification method, we tested the performance of the system for two different frequency decomposition methods (Constant-Q filters and constant bandwidth filters) and two different classification algorithms: linear discriminant analysis (LDA) and probabilistic classification vector machines (PCVMs). We previously found the constant-Q filters can outperform constant bandwidth filters [16] and that PCVM can outperform other commonly used classifier such as LDA and SVM [17]. The following sub sections provide details of each block in in the BCI system.

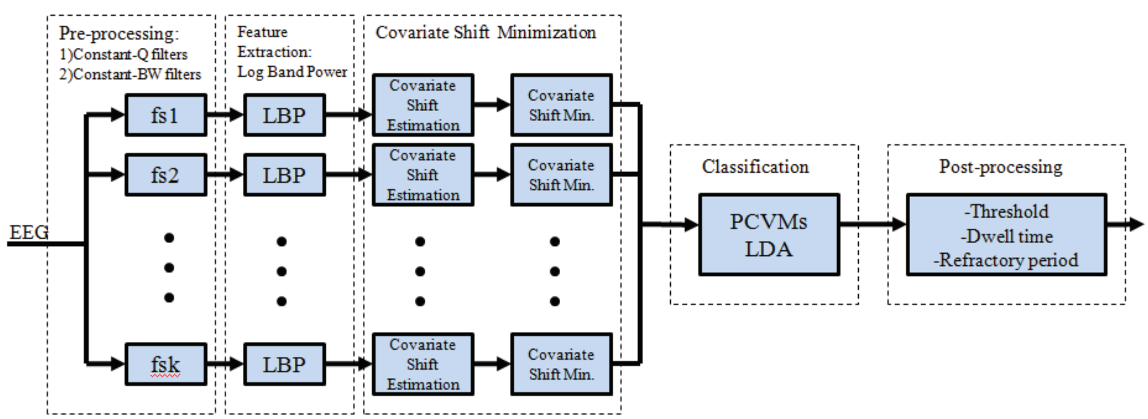


Figure 2. Block diagram of the proposed system where the shift is removed for each feature independently

A. *Frequency decomposition and Feature extraction*

Due to the movement or imagination of the movement, EEG signal energy in specific frequency bands and also in specific brain regions fluctuate, producing an event related desynchronization (ERD) before and during movement and event related synchronization (ERS) in the beta frequency band after termination of the movement [18]. The results of [15] demonstrate that the ERS patterns are more successful in detecting the foot movement (event or control state) in the ongoing (continuous) EEG signal (which contain event as well as non-event information i.e., an idle state).

In this paper we therefore only consider ERS as a neurological phenomenon for discriminating foot movement from the idle state. For detecting ERS from continuous EEG the logarithmic power of the signal in different frequency bands are calculated. In the preprocessing step the signal is filtered using two different filtering methods: 1) constant bandwidth filter bank which gives frequency components from 6 to 36Hz with a length of 2Hz and an overlap of 1 Hz; 2) constant-Q filter bank with 14 center frequencies from 6 to 36Hz and for two different Q values Q= 2 and Q=3[14]. For constant-Q frequency decomposition, the ratio of center frequency to bandwidth for all the filters is the same and equal to Q.

The logarithmic band power features are extracted from time segments of 1s lengths of EEG signal as follows: (i) band-pass filtering using one of the above filtering methods, (ii) squaring the value of each sample, (iii) averaging all samples within the time segment and (iv) applying the logarithm function. Two adjacent segments have an overlap of 0.8 seconds.

The continuous EEG data is categorized into two classes: baseline and movement. All the samples were labeled for the classification of ERS against all other brain activities. According to the ERD/S map of the training runs of the subjects, (Fig. 3) the ERS happens mostly in t=4-5 seconds in each trial. Therefore the samples in t=4-5s of each are labeled as movement/event class (class 2) and the rest of the samples are labeled as baseline/idle (class 1).

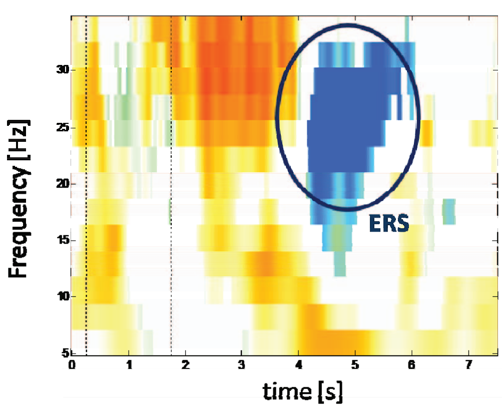


Figure 3. Time-frequency ERD/S map of one subject (s4)

B. *Covariate shift minimization*

As outlined above, one notable representation of nonstationarity in EEG is that feature distributions differ between the training data and test data and this alteration and shift in distribution affects the classification accuracy. In the test phase after extracting logarithmic band power of the signal in different frequencies we estimate the shifts of each feature separately and adapt the features to minimize these shifts. The method is the same as CSM [9] which is the online estimation of the shift occurring in the distributions by means of a polynomial fitting to a short history of feature values.

Suppose we have N consequent 1-sec time segments in the test session (two consequent segments have 0.8s overlap). From each segment we extract a feature vector $\mathbf{f} = \left[f_1 \dots f_k\right]^T$. For each of the feature vector's element $f_i \,(i = 1,\dots,k)$ the polynomial fitting is done separately (in [9] only one feature is adapted). For the first T-1 segments, the extracted features remain unchanged i.e., no adaptation is involved. A polynomial of order h is fitted on these T-1 data points $(x, y_x)\;\; x = 1,\dots,T-1$ with equation (1):

$$y = a_0 + a_1 x + a_2 x^2 + \dots + a_h x^h \tag{1}$$

where $y = f_i$, x is the segment number and $a_0,\dots,a_h$ are polynomial coefficients. In matrix notation, the equation for a polynomial fit is given by $\mathbf{Y} = \mathbf{Xa}$ and this matrix equation can be solved by $\mathbf{a} = \left(\mathbf{X}^{\mathbf{T}}.\mathbf{X}\right)^{-1}.\mathbf{X}^{\mathbf{T}}.\mathbf{Y}$. The polynomial value at the current time segment T using the above equation is calculated and is equal to $\hat{f}_{i(T)}$. The difference between the i-th feature value for the respective time segment ($f_{i(T)}$) and the calculated value using the above polynomial ($\hat{f}_{i(T)}$) shows the i-th feature shift in the time segment T: $shift_{i(T)} = f_{i(T)} - \hat{f}_{i(T)}$. Adding the calculated shift to the common mean of the training feature distribution, $\mu_{0i} = \frac{1}{M}\sum_{j=1}^{M} f_{ji}$ where M is the number of training data, gives the readjusted feature for Tth time segment:

$$Readjusted\ feature_{\,i(T)} = f_{i(T)} - \hat{f}_{i(T)} + \mu_{0i} \tag{2}$$

The coefficients vector $\mathbf{a}$ is continuously updated over T previous time segments and the above calculations are performed for each feature the feature vector (i=1,…,k) and for all time segments from *(T,…,N)* to readjust the features and minimize their shift. Classification of the test features is performed after covariate shift minimization and based on the new adapted features. Selecting the order for polynomial, h, and also the number of previous samples, T-1, for recalculating the polynomial coefficients is one of the challenges of this method. In this paper, the number of previous samples for recalculating the polynomial coefficients and the polynomial order where selected as follows. For different values of T from *30* to *200* samples with step of *20* samples and also for h=1, h=2 and h=3 the performance of the system was calculated using the train data. The best results

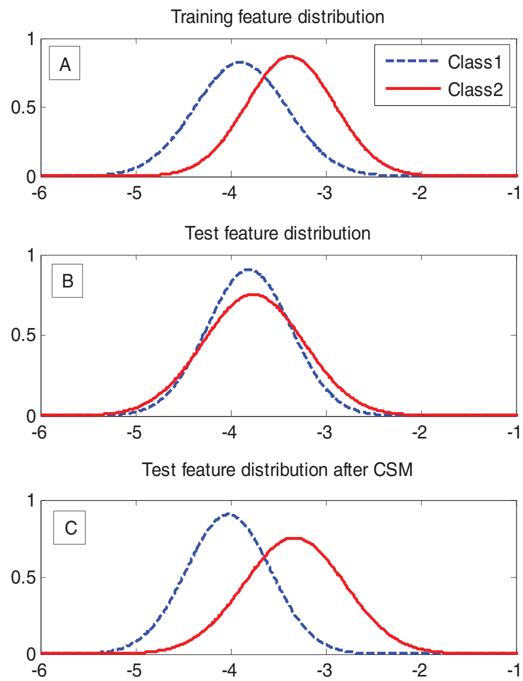


Figure 4. Feature distrubution of A) training data, B) test data and C) test data after applying Covariate shift minimization method for subject S2

were achieved by selecting *T=50* and *h=1* hence these were the values used for the test data. Each feature in the test phase is updated using the features of last 10 seconds. Our system is causal and also is unsupervised since we do not need the information of true class labels of each feature in adaptation process.

An example of applying the CSM on one of the features of the feature vector for subject s2 is plotted in Fig. 4. The first and second rows show the feature distribution of the training and test data, respectively. The shift and change in feature distribution from training to test data is completely clear. In the third row of Fig. 4 the test feature distribution is plotted after features are adapted using CSM.

C. Classification

Two different classification methods, LDA and PCVM, are also compared in this paper, to classify the foot movement or control-events from the baseline resting state EEG signal.

1) Linear Discriminant Analysis (LDA)

The aim of LDA is to define a decision boundary for separating the different classes. The decision boundary is obtained by searching the projection that maximize the distance between the two class means and minimize the interclass variance. The computational requirement of this technique is very low and makes it very suitable for online BCI. Due LDA's simplicity and success in a significant number of motor imagery based BCIs it also applied and tested in the system presented here [19].

2) Probabilistic Classification Vector Machines (PCVMs)

PCVMs is a new sparse learning algorithm proposed in [20] by Chen et al. This classifier not only addresses several drawbacks of support vector machines (SVMs) and also relevant vector machines (RVMs) but also provides some advantages such as: producing probabilistic output for new test points, applying expectation maximization (EM) algorithm instead of computationally expensive grid search by cross validation which is applied in SVMs for finding optimum kernel parameters, and reducing the computational complexity in the test phase as a result of the sparseness-inducing prior of weight vector. Application of this classifier showed performance improvement compared to SVM in detecting foot movement previously [17].

In PCVMs, based on the training set $\{X_i, y_i\}_{i=1}^{N}$, we try to choose a learning model $f(X;W)=\sum_{i=1}^{N} w_i\phi_{i,\boldsymbol{\theta}}(X)+b=\Phi_{\boldsymbol{\theta}}(X)W+b$ where the prediction $f(X;W)$ is a linear combination of N basis functions $\{\phi_{1,\boldsymbol{\theta}}(X),\ldots,\phi_{N,\boldsymbol{\theta}}(X)\}$, (wherein θ is the parameter vector of the basis function), $W=(w_1,\ldots,w_N)^T$ is a parameter of the model and b is the bias. In this paper we used the radial basis function (RBF) as the basis function. In order to specify the model parameters such as W, bias b, and kernel parameters θ, the expectation maximization (EM) algorithm is used. The comprehensive theoretical detail of this classifier is available in [20].

D. Post-processing

The output of the classifiers is post-processed with three simple parameters: one threshold, a dwell time and a refractory period [21]. The amount of time that the output signal of the classifier must exceed the threshold to be considered as a control event is referred to as dwell time. Refractory period is the time interval in which the output signal will be ignored following one control event detection. Using these post-processing parameters limits the false detections of foot movement. Selection of the dwell time and refractory period should be done in a way that allows the system to make fast decisions and also limit the number of detection in each control interval. In this paper the dwell time and the refractory period are fixed for all subjects and are equal to 0.4s and 3s, respectively [17]. Over different threshold values the receiver operating characteristics (ROC) have been analyzed.

IV. RESULTS AND DISCUSSION

In results provided here, the first and second run of data was considered as the train set while third run served as the test set for final evaluations. For evaluation the time interval from t=3 to 5.5 seconds of each trial is considered as the intentional control (IC) period. Performance measurement of the system is carried out in an event by event manner. The event by event TPR and FPR are computed as follows [11]:

$$TPR = \frac{TP}{NTP} \tag{3}$$

$$FPR = \frac{FP}{NFP} \tag{4}$$

$$NFP = \frac{total\ number\ of\ samples\ in\ test\ run}{dwell\ time + Refractory\ period} \tag{5}$$

where *TP* and *FP* are the number of true positive and false positive detections, respectively. *NTP* is the number of IC periods. Table 1 and Table 2 compare the results of foot movement detection (TPR and FPR) with and without covariate shift minimization when two different kinds of filter (constant bandwidth and constant-Q) for preprocessing and two classifiers (LDA and PCVMs) are applied in separate system tests. According to the results of table 1 and 2 the performance of the system the system for all filtering and classifier applied is improved by adapting the features using the CSM (Table 2) compared to no feature adaptation. A two-sided Wilcoxon signed rank between the accuracies obtained with CSM and without CSM showed a significant increase in TPR (p<0.03), and decrease in FPR (p<0.05). Importantly this performance improvement is not just limited to one specific classifier or filtering method, all the columns of table 2 (with CSM) have lower average FPR and higher average TPR compared to the relevant columns in table 1 (without CSM). According to these tables the amount of the performance improvement using CSM is different among subjects since the shifts of test data feature distribution are not equal among subjects. For s2 the improvement after applying CSM is very high while this improvement for s7 is not substantial. The results that the short-term non-stationarity is more evident in EEG signals recorded from subjects s2 and s5.

In this study an unsupervised adaptive method, covariate shift minimization (CSM), has been tested for feature adaptation in a self-paced BCI for the first time. The results show that CSM can track and minimize short-term changes in the feature distribution in an online manner. The major advantage of this method is that it does not require any supervision during the adaptation process. The disadvantage of this approach is that it may not cope well with imbalanced data which could be a major issue in self-paced BCI systems where the idle class generates most of the training/testing samples. For this study the dataset were recorded in a trial based manner but for the real self-paced situation we might have idle state for a long time.

TABLE I. ROC CURVE ANALYSIS OF INDIVIDUAL PERFORMANCE WITHOUT COVARIATE SHIFT MINIMIZATION [%]

Sub[a]	Constant-Bandwidth filters				Constant-Q filters			
	LDA		PCVMs		LDA		PCVMs	
	TPR	*FPR*	*TPR*	*FPR*	*TPR*	*FPR*	*TPR*	*FPR*
S1	100	0	100	0	100	0	100	0
S2	93.3	6.8	60	9.7	76.7	5.5	73	9.7
S3	100	1.3	100	2.7	100	1.4	96.6	1.4
S4	83	5.6	93.3	5.6	93.3	5.6	93.3	2.8
S5	80	18	80	8.3	86.6	8.3	77	8.4
S6	70	9.8	83	9.7	96.6	7	93.3	6.8
S7	93	10	90	4.2	96.6	1.4	96.6	8.4
Avrage	88.4	7	86.6	5.6	92.8	4.2	89.9	5.3

a. subjects

TABLE II. ROC CURVE ANALYSIS OF INDIVIDUAL PERFORMANCE WITH **COVARIATE SHIFT MINIMIZATION** [%]

Sub[a]	Constant-Bandwidth filters				Constant-Q filters			
	LDA		PCVMs		LDA		PCVMs	
	TPR	*FPR*	*TPR*	*FPR*	*TPR*	*FPR*	*TPR*	*FPR*
S1	100	0	100	1.4	100	0	100	0
S2	96.6	5.6	90	6.8	96.6	2.7	96.6	6.8
S3	100	0	100	0.0	100	0	100	1.4
S4	90.0	8.4	97	8.4	96.6	2.8	96.6	4.2
S5	80.0	16.8	90	4.2	90	9.7	90	8.4
S6	93.3	5.6	83	6.8	100	2.8	100	4.2
S7	96.6	3.0	90	6.8	93.3	4.2	96.6	2.8
Avrage	**94.1**	**5.7**	**92.8**	**4.9**	**96.6**	**3.1**	**97.1**	**3.9**

a. subjects

V. CONCLUSION

This study shows for the first time that the performance of the self-paced BCI system in detecting foot movement in continuous EEG is improved by applying the unsupervised covariate shift minimization (CSM) method. In this research, CSM reduces the non-stationary effects of the EEG signal in single session while previous reports show across session improvement in a synchronous BCI paradigm [9]. Future work will include studying the performance of applying CSM method in real-time online BCI systems with feedback to determine the performance improvement given by improved system robustness in the face of non-stationary changes in the continuous EEG signal, both within and across sessions.

ACKNOWLEDGMENT

The authors are grateful to Prof. G. Pfurtscheller and Mr. T. Solis-Escalante of the laboratory of Brain Computer Interface (BCI-Lab), Graz University of technology for making their data available.

REFERENCES

[1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan. "Brain-computer interfaces for communication and control. Clinical neurophysiology," vol.113, no.6, pp.767-791, 2002.

[2] D. J. Krusienski, M. Grosse-Wentrup, F. Galan, D. Coyle, K. J. Miller, E. Forney, and C. W. Anderson, "Critical Issues in Brain Computer Interface Research",*Journal of Neural Eng.*, vol. 8, 025002 (8pp), 2011.

[3] A. Schlögl, C. Vidaurre, K. R. Müller, "Adaptive Methods in BCI Research – An Introductory Tutorial," Brain-Computer Interfaces, Springer, The Frontiers Collection, 2010.

[4] J. Gan, "Self-adapting BCI Based on Unsupervised Learning," in 3rd International Workshop on Brain-Computer Interfaces, Graz, Austria, pp. 50–51, 2006

[5] D.J. McFarland and J. R. Wolpaw, "Sensorimotor rhythm-based brain–computer interface (BCI): feature selection by regression improves performance," *IEEE Trans.Neural Syst. Rehabil. Eng. Vol.*13, pp. 372–9 2005.

[6] A. Satti, D. Coyle and G. Prasad "Continuous EEG classification for a self-paced BCI," *Proc. of the 4th IEEE EMB Conf. on Neural Engineering (Antalya, Turkey)* ,pp.315-8, 2009.

[7] P. Shenoy , M. Krauledat, B. Blankertz , R. P. Rao and K.R. Muller, "Towards adaptive classification for BCI ,"*J. Neural Eng.*vol.**3** pp. 13–23, 2006.

[8] C. Vidaurre, M. Kawanabe, P. von B¨unau, B. Blankertz, and K-R. M¨uller, "Toward an unsupervised adaptation of LDA for Brain-Computer Interfaces," *IEEE Trans. Biomed. Eng*., Vol. 58, no.3, pp.587 -597, 2011

[9] A. Satti , C. Guan, G. Prasad and D. Coyle " A covariate shift minimisation method to alleviate non-stationarity effects for an adaptive brain–computer interface," *20th Int. Conf. Pattern Recognition, PP.105-8,2010.*

[10] M. Sugiyama, M. Krauledat and K.-R. Muller, "Covariate shift adaptation by importance weighted cross validation," *Journal of Machine Learning Research,* vol.8, pp. 985-1005, 2007

[11] Y. Li, H. Kambara, Y. Koike, and M. Sugiyama, "Application of covariate shift adaptation techniques in brain-computer interfaces," *IEEE Trans Biomed. Eng.*, vol. 57, no. 6, pp. 1318–24, 2010

[12] B. Awwad Shiekh Hasan and J.Q. Gan, "Hangman BCI: An unsupervised adaptive self-paced brain-computer interface for playing games" *Comput Biol Med.* Vol.42, no. 5, pp: 598-606, 2012.

[13] J. Millan, and J. Mourino, "Asynchronous BCI and local neural classifiers: an overview of the adaptive brain interface project." *IEEE Trans. Neural Syst. Rehabil. Eng*.,Vol.11, pp.159-161, 2003.

[14] Tsui, C., Gan, J. & Roberts, S. "A self-paced brain-computer interface for controlling a robot simulator: An online event labelling paradigm and an extended kalman filter based algorithm for online training." *Med. & Biol.Eng. Comp.* vol. 47, pp:257:265, 2009

[15] T. Solis-Escalante, G.R. Muller-Putz, G. Pfurtscheller, " Overt foot movement detection in one single Laplacian EEG derivation,"*J. Neurosci. Methods,* Vol. 175, pp.148-53, 2008.

[16] R. Mohammadi, A. Mahloojifar, and D. Coyle "A combination of simple pre and post processing techniques to enhance self-paced BCIs," Advances in Human-Computer interaction, Hindawi press, 185320, (10 pages), 2012.

[17] R. Mohammadi, A. Mahloojifar, H. Chen, and D. Coyle "EEG based Foot Movement Onset Detection with the Probabilistic Classification Vector Machine" *Lecture Notes in Computer Science (LNCS),* ICONIP 2012, vol. 7666, pp 356-363, 2012.

[18] G. Pfurtscheller, and F. H. Lopes da Silva, "Event-related EEG/MEG synchronization and desynchronization: basic principles." *Clin. Neurophysiol,.* Vol. 110, pp. 1842-1857,1999.

[19] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, B. Arnaldi, "*A Review of Classification Algorithms for EEG-based Brain-Computer Interfaces*", Journal of Neural Engineering, 4, R1-R13, 2007.

[20] Chen, H., Tiˇno, P., Yao, X.: Probabilistic Classification Vector Machines. IEEE Trans. Neural Net. 20, 901-14 2009.

[21] Townsend, G., Graimann, B., Pfurtscheller, G.: Continuous EEG classification during motor imagery-simulation of an asynchronous BCI. *IEEE Trans. Neural Syst. Rehabil. Eng*, Vol.12, pp.258-65,2004.